甘蔗 *ScNRAMP* 基因家族的全基因组鉴定与生物信息学分析

刘营[1,2]，尹泽[1,2]，江姚兰[1,2]，周定港[1,2,3*]

（1. 湖南科技大学 生命科学学院，湖南 湘潭 411201；2. 经济作物遗传改良与综合利用湖南省重点实验室，湖南 湘潭 411201；3. 福建农林大学农业部福建甘蔗生物学与遗传育种重点实验室，福州 350002）

摘要：NRAMP 蛋白（natural resistance-associated macrophage proteins）家族在植物响应重金属胁迫时起重要作用，能够转运 Fe^{2+}、Mn^{2+}、Zn^{2+}和 Cd^{2+}等重金属离子。为探究甘蔗 *ScNRAMP* 基因家族的特征，该文基于甘蔗割手密基因组鉴定了 *ScNRAMP* 基因家族，并进行了理化特性、基因结构、顺式作用元件、保守基序、结构域和进化关系等分析。结果表明：甘蔗 *ScNRAMP* 基因家族含有 29 个成员，不均匀分布在 19 条染色体上；编码蛋白均为不稳定蛋白，无信号肽，亚细胞均定位在质膜上；各成员保守基序有 6~10 个不等，跨膜数有 6~12 个不等，二级结构主要构成元件为 α-螺旋和无规则卷曲；顺式作用元件分析表明甘蔗 *ScNRAMP* 基因家族可能通过激素代谢参与逆境胁迫和生长发育等生物过程；利用割手密的 RNA-seq 转录组表达数据进行了组织特异性分析，发现 *ScNRAMP* 在甘蔗不同发育阶段的叶和茎中具有时空表达特性；进化树分析将甘蔗 *ScNRAMP* 家族成员分为 3 个亚家族（I，II 和 III）。该研究在全基因组水平上系统地鉴定了现代栽培甘蔗祖先种之一割手密 *NRAMP* 基因家族，为进一步了解甘蔗 *NRAMP* 基因提供了基础，也为后续甘蔗重金属研究提供了重要候选基因。

关键词：甘蔗，割手密，*ScNRAMP*，基因家族，重金属胁迫

中图分类号：Q943　　　　　**文献标识码：**A

Identification and bioinformatics analysis of *ScNRAMP* gene family in sugarcane

LIU Ying[1,2], YIN Ze[1,2], JIANG Yaolan[1,2], ZHOU Dinggang[1,2,3*]

（1. *School of Life Science*, *Hunan University of Science and Technology*, Xiangtan 411201, Hunan, China; 2. *Hunan Key Laboratory of Economic Crops Genetic Improvement and Integrated Utilization*, Xiangtan 411201, Hunan, China; 3. *Key Laboratory of Sugarcane Biology and Genetic Breeding*, *Ministry of Agriculture*, *Fujian Agriculture and Forestry University*, Fuzhou 350002, China）

Abstract: NRAMP (natural resistance-associated macrophage proteins), which can transport metal ions such as Fe^{2+}, Mn^{2+}, Zn^{2+} and Cd^{2+}, plays essential role in response to heavy metal stress in plant. To better understand the characteristics of the *ScNRAMP* gene family, bioinformatics methods were employed to identify and comprehensively analyze *ScNRAMP* gene family which

---

基金项目：国家自然科学基金青年科学基金项目(31701491)；湖南省自然科学基金（2020JJ50176）；农业部福建甘蔗生物学与遗传育种重点实验室/国家甘蔗工程中心开放课题项目(2017.6.2)[Supported by the National Natural Science Foundation of China(31701491); Natural Science Foundation of Hunan Province, China(2020JJ50176); Key Laboratory of Sugarcane Biology and Genetic Breeding Ministry of Agriculture/ the Open Funding of National Engineering Research Center for Sugarcane, Fujian Agriculture and Forestry University(2017.6.2)]。

作者简介：刘营(1998-)，硕士研究生，研究方向为经济作物遗传育种与栽培，(E-mail)ly1071457534@163.com。

[*] **通信作者：**周定港，博士，副教授，研究方向为植物遗传育种及抗逆改良，(E-mail)dgzhoucn@hnust.edu.cn。

includes protein physicochemical properties, gene structure, conserved motif, domain, cisacting elements and evolutionary relationships. The results were as follows: A total of 29 *ScNRAMP* genes were identified in the *Saccharum spontaneum* genome; These 29 genes were unevenly distributed on 19 chromosomes, and contained 6 to 10 conserved motifs; The encoded proteins are all unstable proteins with no signal peptides and the subcellular locations are all on the plasma membrane; The number of membranes ranges from 6 to 12, and the secondary structure is composed of α-helix and random coils as the main components; Moreover, cis-regulatory elements analysis suggested that *ScNRAMP* may be involved in regulation of stress and development by involving in phytohormone metabolism; The tissue-specific analysis were studied using publicly available gene expression data in ten different tissues and at different developmental stages, which showed that spatiotemporal expression of 29 *ScNRAMP* genes in the leaves and stems of sugarcane at different development stages; Phylogenetic analysis showed that the 29 *ScNRAMP* genes could be divided into three subfamilies (I, II and III). This study could be useful to better understand the *ScNRAMP* gene family and provide significant candidate genes that respond to the stress of heavy metals in sugarcane.

Key words: sugarcane, *Saccharum spontaneum*, *ScNRAMP*, gene family, heavy metal stress

甘蔗(*Saccharum* spp.)是重要的一年生或多年生糖料作物，属单子叶植物纲(Monocotyledoneae)，禾本科(Poaceae)，甘蔗属(*Saccharum* L.)(Irvine, 1999)。甘蔗是一种高光合效率的 C4 植物，具有生物量大、CO_2补偿点低、耐干旱、适应性广和产量高等特点(方静平等, 2014)。甘蔗是异源多倍体，其倍性水平从 5×到 16×不等，其基因组大小约为 10 Gb(Chen et al., 2011)。现代栽培甘蔗是世界上最重要的食糖来源，占全世界食糖总量的 80%(Liu et al., 2020)。

NRAMP 转运体在离子稳态的维持、尤其是二价金属离子的转运上发挥着重要作用，且该蛋白家族在进化过程中高度保守(Nevo & Nelson, 2006)。哺乳动物 *NRAMP* 基因的氨基酸序列与酵母、水稻和果蝇 *NRAMP* 基因的氨基酸序列相似性分别为 46%、58%和 73%(Belouchi et al., 1995)。在植物吸收重金属的过程中，NRAMP 蛋白等膜转运蛋白发挥着重要作用(Cellier et al., 1995)。Mäser et al. (2001)对拟南芥的 6 个 *NRAMP* 基因进行克隆和系统进化分析发现：6 个家族成员被分为两个亚家族，其中 *AtNRAMP*1 和 *AtNRAMP*6 位第一亚家族；而 *AtNRAMP*2~*AtNRAMP*5 属于第二亚家族。AtNRAMP1 蛋白是模式生物拟南芥中的主要高亲和性锰转运蛋白，定位于质膜，其功能是将重金属转运至细胞质(Cailliatte et al., 2010)。Pottier et al. (2015)研究发现诱导表达 *AtNRAMP*4 的酵母株突变会导致其镉和锌的吸收下降，而对于铁的吸收量没有改变，此研究提示人为诱导 NRAMP 转运蛋白的突变可用于降低动植物对 Cd 的吸收和转运。不同植物的 *NRAMP* 家族成员具有组织表达特异性，如 *AtNRAMP1* 在根中高表达(Castaings et al., 2016)，*AtNRAMP2* 则主要在根表皮和根尖区域表达(Gao et al., 2018)；*OsNRAMP1* 在营养生长阶段在根部高表达，而在生殖生长期时在叶片和茎中高表达(Takahashi et al., 2011)。

甘蔗是重要的经济作物，但与其他植物一样也面临着重金属污染问题。重金属对于植物的生长起着十分重要的作用，一些二价重金属离子 Cu^{2+}、Zn^{2+}、Mn^{2+}等，是植物生长必需的微量元素；但有些重金属的过度吸收（如 Cd^{2+}、Pb^{2+}等）则会造成植物的重金属毒害甚至影响产量(Rosa-Santos et al., 2020)。植物 NRAMP 蛋白的功能研究已在拟南芥(Thomine et al., 2000)、水稻(Luo et al., 2018)、土豆(Tian et al., 2021)、大豆(Qin et al., 2017)和油菜(Meng et al., 2017)等植物中广泛报道，而目前还没有关于甘蔗 NRAMP 蛋白的研究报道。研究基于甘蔗割手密种的全基因组序列，利用生物信息学的方法筛选甘蔗 *ScNRAMP* 基因

家族成员，并对其进行进化树构建、染色体定位、基因结构分析和蛋白特性分析等，为后续该基因家族的研究提供参考。

1 材料与方法

1.1 材料

AtNRAMP 基因家族的蛋白序列来源于拟南芥 TAIR 数据库；甘蔗全基因组和基因注释文件来源于 http://www.life.illinois.edu/ming/downloads/Spontaneum_genome/ (Zhang et al., 2018)。

1.2 方法

1.2.1 甘蔗 *ScNRAMP* 基因家族的鉴定

以拟南芥 NRAMP 蛋白序列作为种子序列，在 TBtools（版本号为 v1.0971，Chen et al., 2020）进行第一次 blastp，在 NCBI 进行第二次 blastp，再结合 SMART（http://smart.embl-heidelbreg.de/ ）对候选基因的氨基酸序列结构域进行鉴定，手动剔除不含 NRAMP 保守结构域的序列，保留下的即为甘蔗 *ScNRAMP* 基因家族成员。

1.2.2 编码蛋白理化特性

利用 ExPASy 在线软件（https://us.expasy.org/tools/protparam.html/ ）对 *ScNRAMP* 基因家族编码蛋白的等电点、分子量、序列长度、总平均疏水性、不稳定系数和脂溶指数等理化性质进行分析预测；利用 SignalP v4.1（http://www.cbs.dtu.dk/services/SignalP-4.1/ ）和 CELLO v2.5（http://cello.life.nctu.edu.tw/ ）分别对 29 条蛋白序列进行信号肽预测和亚细胞定位分析。

1.2.3 蛋白二级结构和跨膜结构分析

利用 SOPMA(https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)分析 NRAMP 蛋白的二级结构；利用 TMHMM Server v2.0（http://www.cbs.dtu.dk/services/TMHMM/）分析该蛋白家族的跨膜结构。

1.2.4 甘蔗 *ScNRAMP* 基因保守基序、结构域及基因结构分析

利用 MEME(https://meme-suite.org/meme/)和 NCBI-CDD（https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi）检测甘蔗 *ScNRAMP* 家族基因中所存在相似度较高的基序（motif）和保守结构域（domain）；利用基因组注释文件，获得家族成员内含子及外显子分布情况；最后，利用 TBtools 软件对其进行可视化分析。

1.2.5 甘蔗 *ScNRAMP* 基因家族顺势作用元件分析

提取 *ScNRAMP* 基因上游 3 000 bp 序列作为甘蔗 *NRAMP* 基因的启动子。利用 PlantCARE（http://bioinformatics.psb.ugent.be/webtools/plantcare/html/ ）对启动子区域顺势作用元件进行分析。

1.2.6 甘蔗 *ScNRAMP* 基因的染色体定位

运用 TBtools 软件，结合甘蔗基因组注释文件和家族成员 IDlist，将结果可视化。

1.2.7 甘蔗 *ScNRAMP* 家族进化树的构建

使用本地软件 MEGAX 完成多序列比对并采用邻接法(Neighbor Joining)构建系统进化树，其中校验参数(bootstrap)设置为 1 000 次，其余均设置为默认参数；利用在线 Evolview 软件(https://evolgenius.info//evolview-v2/#login)对进化树进行绘制编辑。

1.2.8 甘蔗 *ScNRAMP* 基因的表达分析

利用割手密的 RNA-seq 转录组表达数据进行甘蔗 *ScNRAMP* 基因的表达分析。割手密的转录组 RNA-seq 表达谱下载自割手密基因组数据库(Li et al., 2020)。筛选得到 *ScNRAMP* 基因家族成员在不同时期不同组织中的表达量（FPKM 值），用 TBtools 进行聚类，绘制热图。

2 结果与分析

2.1 甘蔗 NRAMP 蛋白理化特性

从甘蔗割手密基因组中共鉴定到 29 个甘蔗 *ScNRAMP* 基因家族成员，其一级结构和理化性质如表 1 所示。结果表明：甘蔗 *ScNRAMP* 基因家族成员的氨基酸总数在 334~1 272 个之间，氨基酸长度平均为 563，分子量维持在 55 000 Da 左右；等电点跨度较大，为 4.77~9.43 之间；家族成员的不稳定系数大部分低于 40%，只有 7 个成员不稳定系数介于 40%~48%之间；家族成员平均亲水系数介于 0.015~0.949 之间；29 个甘蔗 ScNRAMP 蛋白均无信号肽，亚细胞定位分析表明成员均定位在质膜上。

2.2 蛋白二级结构和跨膜结构分析

甘蔗 NRAMP 蛋白二级结构和跨膜结构分析如表 2 所示。结果显示：*ScNRAMP* 基因家族的二级结构主要是由 α-螺旋、无规则卷曲、延伸链和β -转角组成，其中 α-螺旋比例最高，达到 36.71%~64.32%，β-转角所占比例最小，占比仅为 1.62%~5.06%；蛋白家族的二级结构较为整齐，除 Sspon.03G0024310-2B 和 Sspon.03G0024310-1A 是无规则卷曲＞α-螺旋＞延伸链＞β-转角，其余家族成员均为 α-螺旋＞无规则卷曲＞延伸链＞β-转角；29 个甘蔗 NRAMP 家族成员均有跨膜结构，其数量为 6~12 个不等。表明甘蔗 NRAMP 蛋白为跨膜蛋白，可能与其转运重金属离子的功能相适应。

2.3 甘蔗 *ScNRAMP* 基因保守基序、结构域及基因结构分析

对甘蔗 *ScNRAMP* 基因的保守基序、结构域及基因结构分析如图 1 所示。保守基序分析发现 motif 1 基序高度保守，在割手密 29 个 NARAMP 家族成员中均存在；个别家族成员存在两个 motif 9，如 *Sspon.02G002160-2C*、*Sspon.02G0057810-1D*、*Sspon.02G002160-1A* 和 *Sspon.02G0021270-1A* 等；大部分家族成员当中有 9 个保守基序，而成员 *Sspon.02G0029680-1A*、*Sspon.05G0038800-1D* 和 *Sspon.02G0057180-1D* 中最少，仅有 6 个保守基序（图 1 和表 3）。保守基序的差异可能暗示该基因家族成员之间在功能上存在差异。

基因结构分析发现 *ScNRAMP* 基因家族全都含有内含子，且内含子数目差异较大（为 2~11 个不等）；部分基因成员在 5′末端和 3′末端不含非翻译区（UTR）。

保守结构域分析发现，除 *Sspon.02G0057810-1D*、*Sspon.05G0038800-1D*、*Sspon.01G0025840-2B*、*Sspon.01G0061080-1D* 和 *Sspon.01G0025840-1A* 这 5 位成员的保守结构域为 Nramp superfamily，其余 24 位家族成员的结构域均为 Nramp。

表 1 甘蔗NRAMP蛋白理化特性

Table 1 Protein physical and chemical properties of NRAMP in sugarcane

ID	氨基酸数目 Number of amino acid	分子量 Mw (Da)	等电点 pI	不稳定系数 Instability index	脂肪系数 Aliphatic index	亲水性的平均值 GRAVY	亚细胞定位 Subcellular localization
Sspon.02G0029680-2B	547	59 123.60	4.85	31.77	111.35	0.459	质膜 Plasma membrane
Sspon.01G0005860-2B	516	56 346.08	5.74	34.63	113.99	0.541	质膜 plasma membrane
Sspon.02G0029690-2D	588	64 938.64	9.08	41.70	107.04	0.270	质膜 Plasma membrane
Sspon.02G0029680-3C	536	57 978.41	4.83	33.99	113.26	0.472	质膜 Plasma membrane
Sspon.02G0029690-1A	538	58 034.37	4.85	35.55	111.78	0.468	质膜 Plasma membrane
Sspon.01G0005860-1P	551	60 825.41	6.28	36.08	111.38	0.464	质膜 Plasma membrane
Sspon.02G0029680-4D	498	53 799.21	4.77	35.72	105.48	0.388	质膜 Plasma membrane

Sspon.01G0005860-1A	507	55 478.08	6.17	33.79	112.76	0.525	质膜 Plasma membrane
Sspon.02G0029680-1A	462	51 417.00	9.59	40.15	103.81	0.147	质膜 Plasma membrane
Sspon.08G0003780-3C	516	55 708.21	8.03	33.01	119.05	0.572	质膜 Plasma membrane
Sspon.02G0021270-1A	548	59 625.14	6.83	38.01	118.87	0.563	质膜 Plasma membrane
Sspon.04G0018970-2B	603	65 379.86	9.05	40.05	108.56	0.483	质膜 Plasma membrane
Sspon.02G0021260-1A	546	58 757.48	6.59	38.89	124.14	0.721	质膜 Plasma membrane
Sspon.08G0003780-1A	490	53 180.59	8.93	31.98	119.96	0.578	质膜 Plasma membrane
Sspon.08G0003780-4D	488	52 797.32	9.34	33.67	120.86	0.631	质膜 Plasma membrane
Sspon.04G0018970-3C	485	52 442.10	8.83	33.01	120.06	0.755	质膜 Plasma membrane
Sspon.02G0021260-2C	514	55 179.28	6.26	37.66	125.97	0.748	质膜 Plasma membrane
Sspon.04G0018970-1A	504	54 945.09	9.43	38.69	116.11	0.612	质膜 Plasma membrane
Sspon.04G0018970-4D	688	73 842.91	9.38	47.12	98.30	0.267	质膜 Plasma membrane
Sspon.02G0021270-2C	450	49 093.10	8.07	34.03	124.80	0.696	质膜 Plasma membrane
Sspon.08G0003780-2B	494	53 497.17	9.19	36.45	124.51	0.646	质膜 Plasma membrane
Sspon.05G0038800-1D	435	46 989.04	6.93	34.97	108.80	0.530	质膜 Plasma membrane
Sspon.03G0024310-3C	717	77 854.58	5.02	37.10	103.07	0.334	质膜 Plasma membrane
Sspon.03G0024310-2B	1 272	138 326.42	5.90	46.56	92.56	0.029	质膜 Plasma membrane
Sspon.02G0057810-1D	308	33 144.69	7.70	41.30	137.40	0.949	质膜 Plasma membrane
Sspon.03G0024310-1A	1 267	138 008.88	5.97	46.07	92.38	0.015	质膜 Plasma membrane
Sspon.01G0025840-2B	456	49 749.03	6.40	31.01	125.68	0.812	质膜 Plasma membrane
Sspon.01G0061080-1D	458	50 027.17	6.05	32.71	124.48	0.742	质膜 Plasma membrane
Sspon.01G0025840-1A	334	36 374.03	6.78	30.42	122.54	0.734	质膜 Plasma membrane

表 2 甘蔗 NRAMP 蛋白二级结构组成及跨膜结构预测结果

Table 2 Secondary structure composition and transmembrane structure prediction of sugarcane NRAMP protein

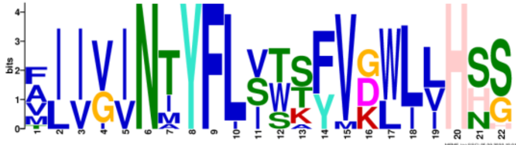
基因 Gene	α-螺旋 α-helix (%)	延伸链 Extended chain (%)	β-转角 β-turn (%)	无规则卷曲 Random coil (%)	跨膜结构数 Number of transmembrane
Sspon.02G0029680-2B	57.22	11.15	3.66	27.97	11
Sspon.01G0005860-2B	57.56	10.08	3.10	29.26	10
Sspon.02G0029690-2D	53.57	12.59	3.06	30.78	11
Sspon.02G0029680-3C	57.65	11.38	2.99	27.99	11
Sspon.02G0029690-1A	58.55	11.15	3.90	26.39	11
Sspon.01G0005860-1P	57.99	11.90	3.35	26.77	8
Sspon.02G0029680-4D	53.61	13.45	3.41	29.52	9
Sspon.01G0005860-1A	57.40	11.05	2.17	29.39	10
Sspon.02G0029680-1A	58.01	13.85	3.03	25.11	6
Sspon.08G0003780-3C	52.91	13.95	3.10	30.04	10
Sspon.02G0021270-1A	53.47	12.23	3.10	31.20	11
Sspon.04G0018970-2B	53.90	12.77	3.65	29.28	10
Sspon.02G0021260-1A	55.31	15.38	2.75	26.56	11
Sspon.08G0003780-1A	54.69	11.84	3.06	30.41	10
Sspon.08G0003780-4D	51.43	13.52	4.52	30.53	10
Sspon.04G0018970-3C	64.32	9.28	2.89	23.51	11
Sspon.02G0021260-2C	55.84	15.95	5.06	23.15	10
Sspon.04G0018970-1A	61.11	12.90	2.18	23.81	9

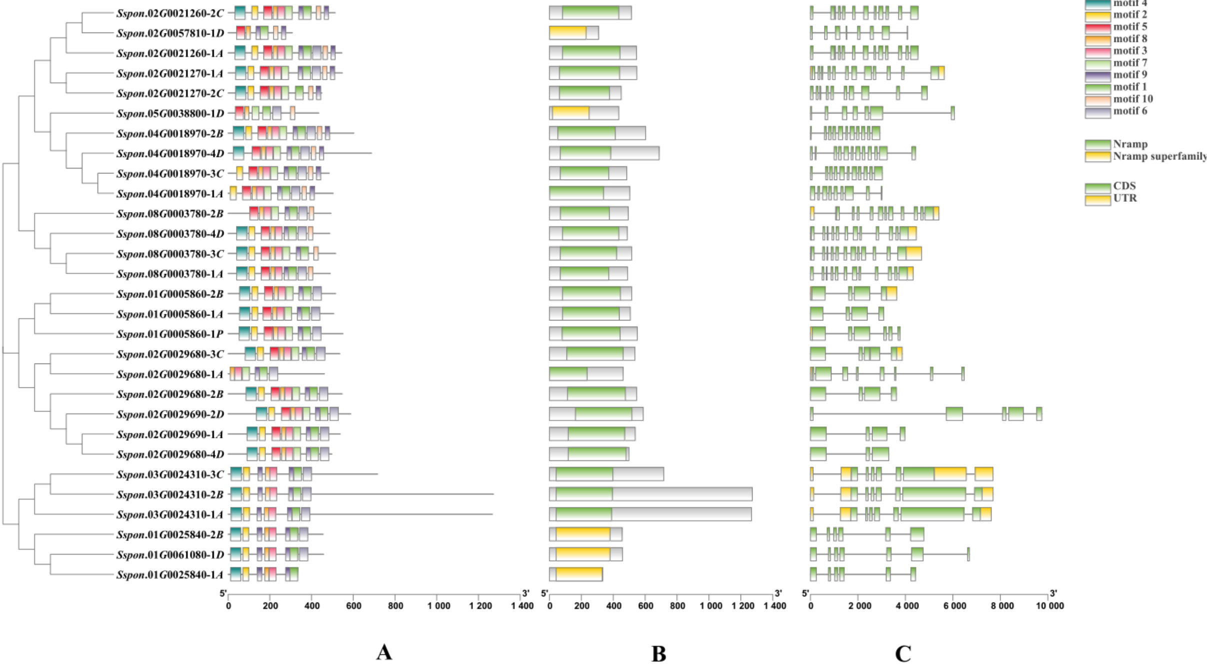
Sspon.04G0018970-4D	43.31	11.92	4.36	40.41	9
Sspon.02G0021270-2C	54.00	15.33	4.67	26.00	9
Sspon.08G0003780-2B	48.58	15.79	1.62	34.01	10
Sspon.05G0038800-1D	46.44	18.16	4.37	31.03	7
Sspon.03G0024310-3C	47.56	13.95	2.65	35.84	12
Sspon.03G0024310-2B	36.71	12.34	2.83	48.11	11
Sspon.02G0057810-1D	53.90	20.78	3.57	21.75	8
Sspon.03G0024310-1A	36.78	12.63	2.92	47.67	11
Sspon.01G0025840-2B	54.17	17.11	3.07	25.66	9
Sspon.01G0061080-1D	53.06	18.78	3.28	24.89	8
Sspon.01G0025840-1A	55.09	19.76	4.19	20.96	6

表 3 甘蔗 *ScNRAMP* 基因保守基序 logo

Table 3 Conservative motif logo of *ScNRAMP* gene in sugarcane

基序 Motif	E-Value	长度 Length	序列标志 Sequence logo
motif 1	6.2e-594	38	
motif 2	3.5e-447	29	
motif 3	3.7e-453	34	
motif 4	4.2e-481	50	
motif 5	3.6e-464	40	
motif 6	42e-399	41	
motif 7	9.4e-306	34	
motif 8	3.8e-229	21	
motif 9	2.2e-206	21	

motif 10	2.5e-129	22	



A. *ScNRAMP* 基因的保守基序（motif）分布，10 个 motif 用不同颜色的方框表示；**B**. *ScNRAMP* 基因保守结构域（domain）分布；**C**. *ScNRAMP* 基因内含子、外显子分布。绿色方框表示外显子；黑色线条表示内含子；黄色框表示的是基因的上下游区域。

A. Distributions of conserved motifs in *ScNRAMP* genes, ten putative motifs are indicated in different colored boxes; **B**. Distributions of domains in *ScNRAMP* genes; **C**. Exon/intron organization of *ScNRAMP* genes. Green boxes represent exons; Black lines with same length represent introns; The upstream/downstream region of *ScNRAMP* genes are indicated in yellow boxes.

图 1 甘蔗 *ScNRAMP* 基因保守基序、结构域及基因结构

Fig.1 Conserved motif, domain and gene structure analysis of *ScNRAMP* gene in sugarcane

2.4 甘蔗 *ScNRAMP* 基因家族顺式作用元件分析

ScNRAMP 基因起始密码子上游 3 000 bp 启动子区域顺式作用元件分析结果如图 2 所示：*ScNRAMP* 基因启动子上含有：（1）与生长发育有关的作用元件：CAT-box（分生组织调控元件）、RY-element（种子特异调控元件）；（2）激素响应相关的顺式作用元件：ABRE（脱落酸响应元件）、TGACG-motif（茉莉酸响应元件）、TCA-element（水杨酸响应元件）TGA-element（生长素响应元件）和 TATC-box/P-box（赤霉素响应元件）；（3）非生物胁迫相关的作用元件：TC-rich（防御和应激响应元件）、MBS（干旱响应元件）、LTR（低温响应元件）、GT1-motif/G-box、MRE 和 I-box（光响应元件）和 ARE（厌氧反应元件）；（4）蛋白相关元件：MBSI 和 MYBHv1（MYB 转录因子结合位点）及 O2-site（玉米醇溶蛋白代谢调节）。进一步分析发现 G-box、MRE、TGACG-motif、ABRE 等元件存在于所有成员。这些结果说明，*ScNRAMP* 基因家族很有可能通过 TC-rich motif 和 MRE motif 等元件参与甘蔗的激素代谢进而影响其对重金属胁迫的响应。

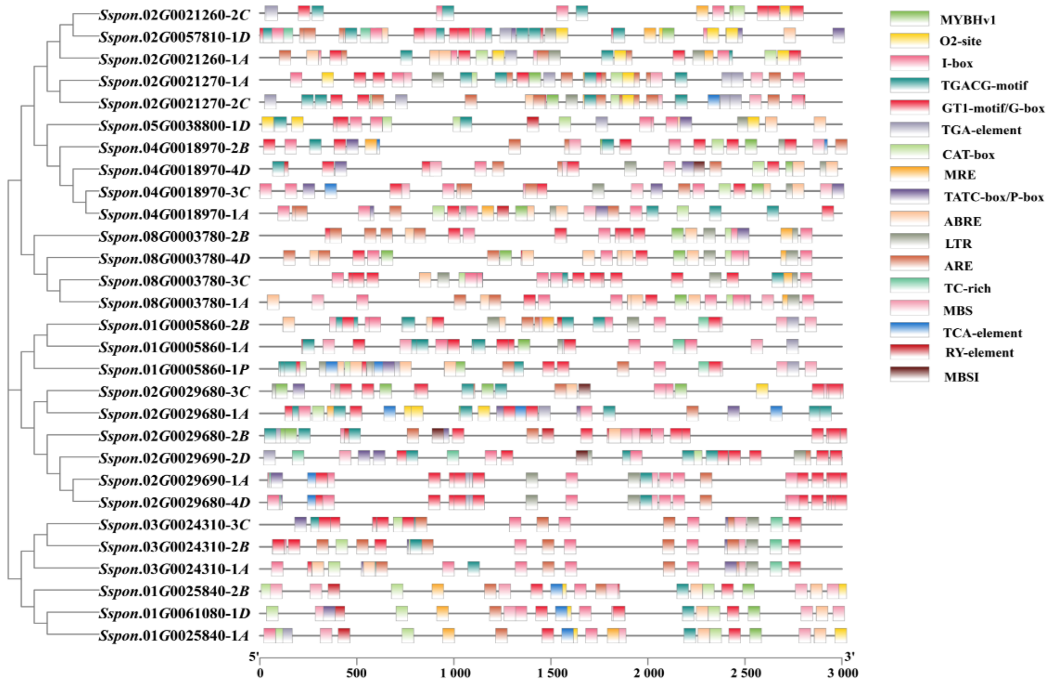


图 2 甘蔗 *ScNRAMP* 基因启动子区域顺势作用元件预测结果

Fig.2 The prediction of cis-elements in the promoters of *ScNRAMP* gene in sugarcane

2.5 甘蔗 *ScNRAMP* 基因的染色体定位

染色体定位结果如图 3 所示。甘蔗 *ScNRAMP* 基因的 29 个家族成员不均一的定位在 19 条染色体上。染色体 Chr2A 上有 4 个基因；染色体 Chr1B、染色体 Chr2C 和染色体 Chr2D 上各有 3 个基因；染色体 Chr1A 上有 2 个基因；其余 14 条染色体上各含有 1 个基因。

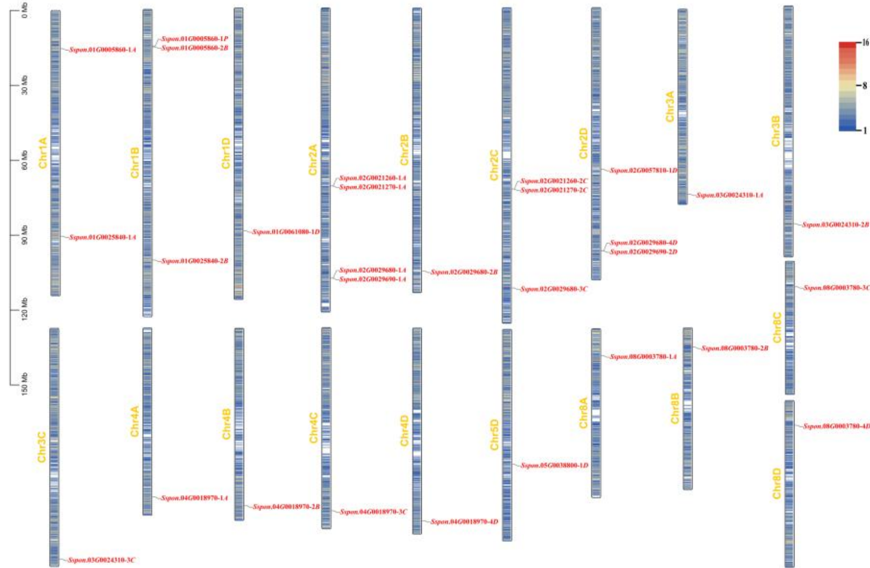
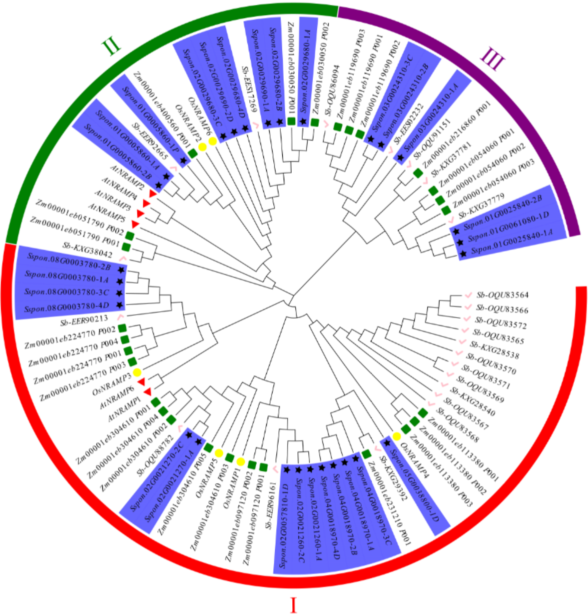


图 3 甘蔗 *ScNRAMP* 基因在染色体上的位置

Fig.3 Position of *ScNRAMP* gene on chromosome in sugarcane

2.6 甘蔗 *ScNRAMP* 家族的系统发育树

为了解甘蔗割手密和其他禾本科植物 *NRAMP* 基因家族的进化关系，构建了水稻、拟南芥、割手密、玉米和高粱五个物种 *NRAMP* 家族成员的进化树（图 4）。进化树有三个分支，分为 3 个亚家族。亚家族 I 中甘蔗 *ScNRAMP* 有 14 位成员（*Sspon.05G0038800-1D*、*Sspon.04G0018970-3C*、*Sspon.04G0018970-1A*、*Sspon.04G0018970-2B*、*Sspon.04G0018970-4D*、*Sspon.02G0021260-1A*、*Sspon.02G0021260-2C*、*Sspon.02G0057810-1D*、*Sspon.02G0021270-1A*、*Sspon.02G0021270-2C*、*Sspon.08G0003780-2B*、*Sspon.08G0003780-1A*、*Sspon.08G0003780-3C* 和 *Sspon.08G0003780-4D*），与高粱 14 位成员、玉米 15 位成员、拟南芥 2 位成员和水稻 4 位成员聚为一支。亚家族 II 中甘蔗 *ScNRAMP* 有 9 位成员（*Sspon.01G0005860-2B*、*Sspon.01G0005860-1A*、*Sspon.02G0029680-2B*、*Sspon.02G0029680-1A*、*Sspon.02G0029680-3C*、*Sspon.02G0029680-4D*、*Sspon.02G0029690-1A*、*Sspon.02G0029690-2D* 和 *Sspon.01G0005860-1P*），与高粱 3 位成员、玉米 5 位成员、拟南芥 4 位成员和水稻 2 位成员聚为一支。亚家族III中甘蔗 *ScNRAMP* 有 6 位成员（*Sspon.03G0024310-3C*、*Sspon.03G0024310-2B*、*Sspon.03G0024310-1A*、*Sspon.01G0025840-2B*、*Sspon.01G0061080-1D* 和 *Sspon.01G0025840-1A*），与高粱 5 位成员和玉米 7 位成员聚为一支，且在此分枝上没有拟南芥和水稻的家族成员。水稻、拟南芥、高粱、玉米和甘蔗 *NRAMP* 家族成员之间均有同源基因，相比之下玉米与甘蔗的 *NRAMP* 基因家族成员之间亲缘关系更近。



黑色星号表示甘蔗 *ScNRAMP*；绿色方框表示玉米 *ZmNRAMP*；黄色圆圈表示水稻 *OsNRAMP*；红色三角形表示拟南芥 *AtNRAMP*；粉色对号表示高粱 *SbNRAMP*。

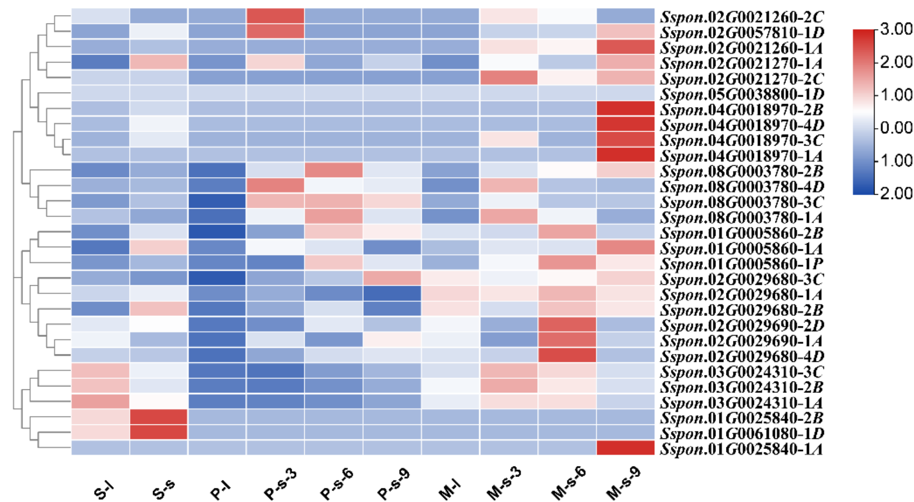
Black asterisks represent *ScNRAMP*; Green boxes represent *ZmNRAMP*; Yellow circles represent *OsNRAMP*; Red triangles represent *AtNRAMP*; Pink ticks represent *SbNRAMP*.

图 4 甘蔗 *ScNRAMP* 的系统发育树

Fig.4 Phylogenetic tree of *ScNRAMP* in sugarcane

2.7 甘蔗 *ScNRAMP* 基因组织表达

为了解甘蔗 *ScNRAMP* 基因家族成员的功能与时空表达模式，我们以公开发表的割手密基因表达数据分析了 *ScNRAMP* 基因在不同组织和不同发育时期的表达情况，如图 5 所示。结果显示：甘蔗 *ScNRAMP* 基因的表达呈现组织特异性；*Sspon.04G0018970-2B*、*Sspon.04G0018970-4D*、*Sspon.04G0018970-1A*、*Sspon.04G0018970-3C* 等基因在茎成熟期（mature-steam-3、6、9）表达量较高，在前成熟期和发芽期表达量较低；*Sspon.03G0024310-3C*、*Sspon.03G0024310-2B*、*Sspon.03G0024310-1A*、*Sspon.01G0025840-2B* 和 *Sspon.01G0061080-1D* 五个成员在发育期叶中的表达量高于前成熟期和成熟期。另外，29 个家族成员中仅有 1 个成员 *Sspon.05G0038800-1D* 的表达量 FPKM 值为 0，表明其在各发育时期和组织部位均未表达。



S 表示实生苗期；P 表示成熟前期；M 表示成熟期；s 表示茎；l 表示叶；3 表示甘蔗茎第 3 节；6 表示甘蔗茎第 6 节；9 表示甘蔗茎第 9 节。

S represents seedling stage; P represents early maturity stage; M represents mature stage; s represents steam; l represents leaf; 3 represents section 3 of sugarcane; 6 represents section 6 of sugarcane; 9 represents section 9 of sugarcane.

图 5 甘蔗 *ScNRAMP* 基因表达模式分析

Fig.5 Expression profile analysis of *ScNRAMP* genes in sugarcane

3 讨论与结论

对甘蔗 *ScNRAMP* 基因家族成员进行理化特性分析发现，不同的 NRAMP 蛋白序列有较大的差异，氨基酸长度为 334~1 272 aa，等电点跨度较大（4.77~9.43），表明其编码蛋白能适应不同的酸碱环境。甘蔗 NRAMP 蛋白不稳定系数均小于 40%，表明其稳定性较好；蛋白的平均亲水系数介于 0.015~0.949 之间，表明其是一类相对疏水的蛋白。29 个家族成员定位在 19 条染色体上，染色体 Chr2A 上最多，有 4 个家族成员。割手密 NRAMP 蛋白所有成员均有 6~12 个跨膜结构域，均定位在质膜上，可能与其具有转运重金属离子的功能相适应。基因结构分析显示所有家族成员均具有内含子。

甘蔗割手密基因组中共鉴定出 29 个 *ScNRAMP* 基因，与水稻（6 个成员）和拟南芥（6 个成员）相比，*ScNRAMP* 基因家族成员的数量明显要多，推测可能与甘蔗是多倍体、且基因组庞大有关，尤其是甘蔗作为多倍体植物，其在多倍化进程中的全基因组复制促进了甘蔗基因组含量的提高及其基因家族的扩张，这与相关研究茄科植物高度变异的基因家族在全基因组复制和基因的串联重复引起基因组的含量增加和基因家族扩张的观点一致(Mäser et al., 2001；Zhang et al., 2018；Tian et al., 2021；Wang et al., 2021)。拟南芥和水稻的基因

组大小分别为 125 Mb 和 466 Mb，而甘蔗割手密的基因组大小达到了 3.36 Gb，此外，拟南芥和水稻均为二倍体，甘蔗为八倍体，而测序的甘蔗品种割手密 AP85-441 是甘蔗八倍体 SES208 单花粉培育得到的整倍体（四倍体），理论上现代甘蔗栽培种（多为异源八倍体、且为非整倍体）具有的 *ScNRAMP* 家族成员应多于 29 个（Chen et al., 2011; Zhang et al., 2018)。

基因结构和基序组成分析可为基因家族的进化关系提供重要依据（Boudet et al., 2001)。*ScNRAMP* 基因结构分析发现同一亚家族的大多数基因在外显子、基序或非翻译区具有类似的结构特征，这一现象与其他物种的 NRAMP 蛋白家族相似(Belouchi et al., 1997；Lanquar et al., 2005)，如 *AtNRAMP3* 和 *AtNRAMP4* 基因结构相似，且对 Fe^{2+}都有极高运输能力(Lanquar et al., 2005)； 而 *OsNramp1*、*OsNramp2* 和 *OsNramp3* 序列的相似性达到了 64 %~75 %(Belouchi et al., 1997)，推测甘蔗 NRAMP 蛋白同一亚家族的成员在功能上可能相似。甘蔗割手密、水稻、高粱和玉米 4 种单子叶植物和拟南芥 1 种双子叶植物的系统进化分析可将 29 个 *ScNRAMP* 家族成员分为 3 个亚家族。在亚家族 III 中，进化树分支上只有高粱（5 位成员）、玉米（7 位成员）和割手密（6 位成员）的家族成员，而没有水稻和拟南芥的成员，推测甘蔗割手密 *NRAMP* 亚家族 III 中的这 6 位成员可能存在着与水稻和拟南芥的 NRAMP 蛋白家族不一样的特殊功能，而这个亚家族的特殊功能在高粱和玉米中也可能存在。

植物启动子作为植物功能基因表达调控的最重要顺式元件之一，启动子分析能够帮助我们阐明基因表达的调控和响应机制。通过对启动子区域的顺式作用元件分析可得到植物关于响应特定生物胁迫或非生物胁迫的重要元件。通过启动子区域的顺式作用元件分析，推测甘蔗割手密种可能通过光响应因子、激素响应元件和抗逆响应因子（TC-rich motif）等元件间的相互作用形成复杂的代谢调控网络来响应和应对重金属胁迫。这与水稻 *NRAMP* 基因受到 JA、ABA 等激素的调控来应对病原菌侵染、外界防御信号和金属离子，从而参与植物的防御反应一致(Zhou et al., 2004)。在对甘蔗 *ScNRAMP* 上游 3 000 bp 区域研究发现成员中大部分均含有茉莉酸响应元件（TGACG-motif）和生长素响应元件（TGA-element），推断该基因家族在进化过程中可能具有与植物生长和逆境胁迫相关的功能。甘蔗 *ScNRAMP* 基因的表达具有组织特异性，甘蔗 *ScNRAMP* 基因主要在特定组织和器官中表达，*Sspon.01G0025840-2B* 和 *Sspon.01G0061080-1D* 两位成员在种子发芽期的叶和茎中的表达量明显高于其他时期。对各时期和各组织高表达的基因分析，发现家族成员均具有 MBS、ABRE 和 TCA-element 等与激素响应相关的作用元件，推测其参与激素代谢过程进而影响甘蔗对重金属胁迫的响应。

目前，*NRAMP* 基因家族在菜豆(Ishida et al., 2018)、大豆(Qin et al., 2017)、甘蓝型油菜(Meng et al., 2017)等经济作物上已经被广泛研究。甘蔗作为重要的糖料作物和能源作物，关系着我国乃至世界人民的食糖安全，降低重金属胁迫对甘蔗产量及含糖量的影响，对 *ScNRAMP* 等基因家族转运重金属离子的分子机制的了解尤其必要。本文通过对甘蔗割手密中 29 个 *ScNRAMP* 基因家族成员的全基因组分析，利于阐明甘蔗 *ScNRAMP* 基因家族的功能和为甘蔗的分子育种尤其是应对重金属胁迫的研究提供重要的候选基因。

参考文献：

BELOUCHI A, CELLIER M, KWAN T, et al., 1995. The macrophage-specific membrane protein Nramp controlling natural resistance to infections in mice has homologues expressed in the root system of plants [J]. Plant Mol Biol, 29: 1181-1196.

BELOUCHI A, KWAN T, GROS P, 1997. Cloning and characterization of the *OsNramp* family from *Oryza sativa*, a new family of membrane proteins possibly implicated in the transport of

metal ions [J]. Plant Mol Biol, 33: 1085-1092.
BOUDET N, AUBOURG S, KREIS M, et al., 2001. Evolution of intron/exon stucture of *DEAD* helicase family genes in *Arabidopsis*, *Caenorhabditit* and *Drosophila* [J]. Genome Res, 11(12): 2101-2114.
CAILLIATTE R, SCHIKORA A, BRIAT JF, et al., 2010. High-affinity manganese uptake by the metal transporter *NRAMP1* is essential for *Arabidopsis* growth in low manganese conditions [J]. Plant Cell, 22: 904-917.
CASTAINGS L, CAQUOT A, LOUBET, et al., 2016. The high-affinity metal transporters *NRAMP1* and *IRT1* team up to take up to iron under sufficient metal provision [J]. Sci Rep, 6: 37222.
CELLIER M, PRIVÉ G, BELOUCHI A, et al., 1995. Nramp defines a family of membrane proteins [J]. Proc Natl Acad Sci USA, 92: 10089-10093.
CHEN CJ, CHEN H, ZHANG Y, et al., 2020. Tbtools: an integrative toolkit developed for interactive analyses of big biological data [J]. Mol Plant, 13: 1194-1202.
CHEN RK, XU LP, LIN YQ, et al., 2011. Modern sugarcane genetic breeding [M]. Beijing: China Agricultural Press: 329-372.
FANG JP, 2014. A review of *Saccharum* origin and its evolutionary relationship with related genera [J]. Chin J Trop Crops, 35(4): 816-822. [方静平, 2014. 甘蔗属起源及其与近缘属进化关系研究进展 [J]. 热带作物学报, 35(4): 816-822.]
IRVINE JE, 1999. *Saccharum* species as horticultural classes [J]. Theor Appl Genet, 98: 186-194.
ISHIDA JK, CALDAS DG, OLIVEIRA LR, et al., 2018. Genome-wide characterization of the *NRAMP* gene family in *Phaseolus vulgaris* provides insights into functional implications during common bean development [J]. Genet Mol Biol, 41: 820-833.
LANQUAR V, LELIÈVRE F, BOLTE S, et al., 2005. Mobilization of vacuolar iron by *AtNRAMP3* and *AtNRAMP4* is essential for seed germination on low iron [J]. EMBO J, 24: 4041-4051.
LI PT, CHAI Z, LIN PP, et al., 2020. Genome-wide identification and expression analysis of *AP2/ERF* transcription factors in sugarcane (*Saccharum spontaneum* L.) [J]. BMC Genomics, 21(1): 685.
LIU XL, YIN Z, LIU Y, et al., 2020. The complete mitochondrial genome of sugarcane (*Saccharum* spp.) variety FN15 [J]. Mitochondrial DNA Part B, 5: 2163-2165.
LUO BB, CHEN JG, ZHU LL, et al., 2018. Overexpression of a high-affinity nitrate transporter *OsNRT2.1* increases yield and manganese accumulation in rice under alternating wet and dry condition [J]. Front Plant Sci, 9: 1192.
MÄSER P, THOMINE S, SCHROEDER JI, et al., 2001. Phylogenetic relationships within cation transporter families of *Arabidopsis* [J]. Plant Physiol, 126: 1646-1667.
MENG JG, ZHANG XD, TAN K, et al., 2017. Genome-wide identification of Cd-responsive *NRAMP* transporter genes and analyzing expression of *NRAMP1* mediated by miR167 in *Brassica napus* [J]. Biometals, 30: 917-931.
NEVO Y, NELSON N, 2006. The NRAMP family of metal-ion transporters [J]. Biochim Biophys Acta, 1763: 609-620.
POTTIER M, OOMEN R, PICCO C, et al., 2015. Identification of mutations allowing Natural Resistance Associated Macrophage Proteins (NRAMP) to discriminate against cadmium [J]. Plant J, 83: 625-637.
QIN L, HAN PP, CHEN LY, et al., 2017. Genome-wide identification and expression analysis of

NRAMP family genes in Soybean (*Glycine Max* L.) [J]. Front Plant Sci, 8: 1436.

ROSA-SANTOS TM, SILVA RG, KUMARP, et al., 2020. Molecular mechanisms underlying sugarcane response to aluminum stress by RNA-Seq [J]. Int J Mol Sci, 21(21): 7934.

TAKAHASHI R, ISHIMARU Y, NAKANISHI H, et al., 2011. Role of the iron transporter *OsNRAMP1* in cadmium uptake and accumulation in rice [J]. Plant Signal Behav, 6(11): 1813-1816.

THOMINE S, WANG R, WARD JM, et al., 2000. Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to *Nramp* genes [J]. Proc Natl Acad Sci USA, 97: 4991-4996.

TIAN WJ, HE GD, QIN LJ, et al., 2021. Genome-wide analysis of the *NRAMP* gene family in potato (*Solanum tuberosum*): Identification, expression analysis and response to five heavy metals stress [J]. Ecotoxicol Environ Saf, 208: 111661.

WANG P, MOORE BM, PANCHY NL, et al., 2018. Factors influencing gene family size variation among related species in a plant family, Solanaceae [J]. Genome Biol Evol, 10(10): 2596-2613.

ZHANG JS, ZHANG XT, TANG HB, et al., 2018. Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L.[J]. Nat Genet, 50: 1565-1573.

ZHOU XJ, YANG YN, 2004. Differential expression of rice *Nramp* genes in response to pathogen infection, defense siganl molecules and metal ions [J]. Mol Plant Pathol, 65(5): 235-243.